Title: Macroevolutionary patterns in overexpression of tyrosine: an anti-herbivore defense in a speciose tropical tree genus, *Inga* (Fabaceae)

Authors: Phyllis D. Coley[1*], María-José Endara[1, 2], Gabrielle Ghabash[1], Catherine A. Kidner[3, 4], James A. Nicholls[5], R. Toby Pennington[4,6], Anthony G. Mills[1], Abrianna J. Soule[1], Maristerra R. Lemes[7] and Thomas A. Kursar[1]

[1]Department of Biology, University of Utah, Salt Lake City, UT, USA

[2]Centro de Investigación de la Biodiversidad y Cambio Climático (BioCamb) e Ingeniería en Biodiversidad y Recursos Genéticos. Facultad de Ciencias de Medio Ambiente, Universidad Tecnológica Indoamérica, Quito, Ecuador

[3]Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh, Edinburgh, UK

[4]Royal Botanic Garden, Edinburgh, UK

[5]Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK; current: CSIRO, Australian National Insect Collection, Black Mountain Labs, Acton, ACT 2614, Australia

[6]Department of Geography, University of Exeter, Exeter, EX4 4QE, UK

[7]Laboratório de Genética e Biologia Reprodutiva de Plantas, Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia, Avenida André Araújo, 2936, Petrópolis, 69067-375, Manaus, AM, Brazil

*corresponding author, coley@biology.utah.edu

key words: plant-herbivore interactions, tyrosine overexpression, tropical rainforests, *Inga* (Fabaceae), expanding leaves, prephenate dehydrogenase, herbivore host choice

Abstract

1. Plant secondary metabolites are a key defense against herbivores, and their evolutionary origin is likely from primary metabolites. Yet for this to occur, an intermediate step of overexpression of primary metabolites would need to confer some advantage to the plant. Here we examine the evolution of overexpression of the essential amino acid, L-tyrosine, and its role as a defense against herbivores.

2. We examined overexpression of tyrosine in 97 species of *Inga* (Fabaceae), a genus of tropical trees, at five sites throughout the Neotropics. We predicted that tyrosine could act as an anti-herbivore defense because concentrations of 4% tyrosine in artificial diets halved larval growth rates. We also collected insect herbivores to determine if tyrosine and its derivatives influenced host associations.

3. Overexpression of tyrosine was only present in a single lineage comprising 21 species, with concentrations ranging from 5-20% of the leaf dry weight. Overexpression was pronounced in

expanding but not in mature leaves. Despite laboratory studies showing toxicity of L-tyrosine, *Inga* species with tyrosine suffered higher levels of herbivory. We therefore hypothesize that overexpression is only favored in species with less effective secondary metabolites. Some tyrosine-producing species also contained secondary metabolites that are derived from tyrosine: tyrosine-gallates, tyramine-gallates and DOPA-gallates. Elevated levels of transcripts of prephenate dehydrogenase, an enzyme in the tyrosine biosynthetic pathway that is insensitive to negative feedback from tyrosine, were found only in species that overexpress tyrosine or related gallates. Different lineages of herbivores showed contrasting responses to the overexpression of tyrosine and its derived secondary metabolites in their host plants.
4. *Synthesis*. We propose that overexpression of some primary metabolites can serve as a chemical defense against herbivores, and are most likely to be selected for in species suffering high herbivory due to less effective secondary metabolites. Overexpression may be the first evolutionary step in the transition to the production of more derived secondary metabolites. Presumably, derived compounds would be more effective and less costly than free tyrosine as anti-herbivore defenses.

Introduction

Secondary metabolites are one of the most effective defenses plants have against insect herbivores, and play a central role in shaping both ecological and evolutionary interactions. For example, the high local diversity of trees in tropical rainforests may be maintained, not because coexisting species differ in their abiotic niches (Wright, 2002, Kraft, Valencia & Ackerly 2008, Sedio, Wright, & Dick 2012), but because they differ in their secondary metabolites, and therefore do not share herbivores (Kursar et al., 2009, Coley & Kursar, 2014, Salazar, Jaramillo, & Marquis, 2016a & b, Vleminckx et al., 2018, Endara et al., 2018). Because there appear to be an infinite number of chemical combinations, this translates into an infinite number of niches with respect to herbivore pressure. In addition, secondary metabolites have been implicated in the co-diversification of both plants and herbivores (Ehrlich & Raven 1964, Marquis et al., 2016, Endara et al., 2017). Unlike traits associated with resource acquisition (Endara et al., 2015, Sedio et al., 2012), secondary metabolites show no or only a weak phylogenetic signal across closely related species, implying that selection by herbivores favors rapid evolutionary divergence among relatives (Becerra 1997; Sedio 2013, Fine et al., 2013a & b, Salazar et al. 2016a & b, Salazar et al. 2018). In contrast, related herbivores feed on host species with similar secondary metabolites, suggesting that counter-adaptations to plant defenses are evolutionarily conserved (Janz, 2011, Endara et al., 2017). Thus, secondary metabolites appear to play an instrumental role in both the maintenance and origin of diversity for tropical plants and herbivores.

Although overexpression of primary metabolites has rarely been reported, high concentrations of these compounds in the leaf may function as a defensive chemical against herbivores, and may be the first step in the evolution of more biosynthetically-derived secondary metabolites (Lokvam, Brenes-Arguedes, Lee, Coley, & Kursar 2006). Previously, we found overexpression of the essential amino acid L-tyrosine in one species, *Inga umbellifera* (Lokvam et al. 2006). Tyrosine is a key nutrient for herbivores, involved in protein biosynthesis, cuticle hardening, and innate immune responses (Vavricka et al., 2014). Tyrosine is generally present at very low concentrations in mature leaves, including mature leaves of *I. umbellifera* (0.05% DW, Lokvam et al., 2006). Thus, slightly elevated tyrosine concentrations can have a positive effect on herbivore growth (Lokvam et al., 2006). However, concentrations in young *I. umbellifera* leaves were ~10% DW, which proved to be highly toxic to a generalist herbivore. When L-

tyrosine purified from *I. umbellifera* was incorporated into artificial agar diets at a concentration of 3.8% dry mass, larval growth of *Heliothis virescens* (Lepidoptera: Noctuidae) was reduced by 50% compared to controls ($p<0.01$). At 10% dry mass of the diet, equivalent to the concentration in the *I. umbellifera* leaves, growth was reduced to 2% of controls ($p<0.001$) and mortality was high (Lokvam et al., 2006). Thus, overexpression of a primary metabolite functioned as a chemical defense in this species.

Although the toxic effects of L-tyrosine are not well studied, evidence suggests that ingestion is also detrimental to other insects. Blood-feeding kissing bugs (*Rhodnius)* and mosquitos (*Aedes)* ingest large amounts of tyrosine, which are lethal if their specialized enzymes for degradation are blocked (Sterkel et al. 2016). High hemolymph tyrosine has been implicated in increased fluid secretion from the Malpighian tubules of both *Rhodnius* (Hemiptera) and *Drosophila* (Hazel, Ianowski, Christensen, Maddrell & O'Donnell, 2003), potentially disrupting an insect's water balance. In rats, tyrosine causes oxidative stress leading to damage of DNA, lipids and proteins (De Prá et al. 2014). DOPA, a derivative of tyrosine found in some *Inga* species, caused severe cuticular deformities for a lepidopteran herbivore (*Spodoptera*) with the addition of only 0.25% dry mass to artificial diets (Rehr, Janzen and Feeney 1973). Thus, ingestions of tyrosine may negatively impact a variety of herbivores.

Primary metabolites serve as the biosynthetic precursors for most secondary metabolites, suggesting that gene duplication of primary metabolites and subsequent changes to regulation must have occurred (Wink 2016). Yet for this evolutionary scenario to be selected for, the intermediate step of overexpression of primary metabolites would need to be neutral or confer some advantage to the plant. This presents a puzzle, as normal concentrations of primary metabolites are attractive to herbivores. Instead, higher concentrations would have to be toxic, and at least as effective as the secondary metabolites synthesized by that species. Furthermore, overexpression may incur high costs if metabolites are in high concentrations. Thus, we propose that overexpression would most likely be selected for in cases where the suite of secondary metabolites was not effective against herbivores. Over evolutionary time, selection should favor biosynthetic derivatives that would be more effective and less costly than the primary metabolites.

To determine the generality, possible adaptive value and the phylogenetic patterns of tyrosine overexpression, we examined numerous species in the Neotropical tree genus *Inga* (Fabaceae). *Inga* has over 300 described species (Pennington, 1997) and appears to have explosively radiated ~9 million years ago (Richardson, Pennington, Pennington, & Hollingsworth, 2001). Furthermore, in any given forest, it is one of the most diverse genera and one of the most abundant in terms of individual plant stems (Valencia et al., 2004, Dexter et al., 2017). For *Inga,* as with most tropical species, herbivores prefer to feed on expanding leaves, as these are low in fiber and high in nutrients - two unavoidable consequences of growth (Coley, Endara, & Kursar 2018). Even though leaves can live for years, and are only expanding for a few weeks, ~70% of the lifetime damage occurs during this short window between emerging from the bud and reaching full size. As a consequence, there is strong selection to invest in chemical defenses, and in *Inga*, secondary metabolites comprise over 50% of the dry weight of expanding leaves. Once the leaf is full size and can toughen, investment drops by more than half (Wiggins, Forrister, Endara, Coley, & Kursar 2016).

To test the adaptive significance and macroevolutionary patterns of tyrosine overexpression we address the following questions using *Inga* as a model system: 1) Could overexpression of a primary metabolite function as a defense? 2) Why would it evolve? 3) What

is the evolutionary history of tyrosine overexpression and production of derived secondary metabolites? 4) What biosynthetic innovations might allow overexpression? 5) Does overexpression shape host choice by herbivores?

Materials and Methods

Study sites

We sampled *Inga* at five lowland tropical rainforest sites spread across its range (Fig. S1) between 2005 and 2014: Panama (Smithsonian Tropical Research Institute on Barro Colorado Island, 9°N, 80°W), French Guiana (Nouragues Station, 4°N, 53°W), Brazil (Biological Dynamics of Forest Fragments, km 41, 2°S 60°W), Peru (Los Amigos Biological Station, Madre de Dios, 13°S, 70°W) and Ecuador (Tiputini Biodiversity Station, 0.6°S. 76°W). We spent approximately 16 people-months per site collecting data in the field.

Study species and phylogeny

At each site, we collected data on saplings for all of the most common species of *Inga.* Species were identified in the field using vegetative characteristics by Kursar and Coley, who have been studying *Inga* for over 15 years. Leaf samples of representative individuals of each species were dried in silica and later extracted and DNA sequenced to confirm species delineations and to place species in a phylogenetic context. Previous phylogenetic work on *Inga*, using 6000+ base pairs from seven chloroplast DNA loci and one nuclear locus (nuclear ribosomal internal transcribed spacers), generated through Sanger sequencing, resulted in a poorly resolved phylogeny with multiple polytomies (Richardson et al. 2001, Kursar et al. 2009, Dexter et al. 2017). We therefore employed transcriptome data from three *Inga* species to design a bait-capture set that targeted 264 nuclear loci for enrichment from whole-genome libraries for subsequent sequencing (Nicholls *et al.,* 2015). A preliminary unpublished maximum-likelihood analysis of these hybrid-capture data (over 300,000 base pairs) showed excellent resolution of the *Inga* phylogeny, with high support for nearly all nodes at all depths within the phylogeny. It provided clear resolution of relationships among major clades as well as among closely related *Inga* species and populations within species. We also used untargeted metabolomic fingerprints (chemocoding) of species and populations to further confirm our classifications (Endara et al., 2018).

Leaf collections

To determine tyrosine levels, we collected expanding and mature leaves from understory saplings between 0.5m and 3m tall. Expanding leaves were between 20 and 90% of full size and were tender. Mature leaves were tough and, based on color, were 3-6 months post expansion. We made collections of expanding leaves from 97 species of *Inga*. We made collections of mature leaves from a subset of these species, as mature leaves were not the original focus of the project and we were unable to recollect all species. Because many species occurred at several sites, we sampled 159 species/site combinations. Several individuals of each species at each site were collected for a total of 569 plants. If a species overexpressed tyrosine it was sampled at all the sites where it occurred.

Chemical analyses

In the field, leaf collections were dried in silica at ambient temperature, and then sent to Utah and stored at -80ºC until analysis. Because tyrosine is not soluble in standard methanolic

extractions, leaf samples (5mg) were pulverized and extracted for 15 minutes in 2 mL 0.5% acetic acid at 80ºC, centrifuged twice, and analyzed by reverse-phase High Performance Liquid Chromatography (RP-HPLC) using a diode array detector (DAD) at 274 nm (see Supporting Information for details). In addition to tyrosine, this method could also detect phenylalanine and tryptophan, although these two amino acids were not observed in any leaf samples.

Ultra Performance Liquid Chromatography-Mass Spectrometry (UPLC-MS) was also used to confirm that the L-tyrosine detected using HPLC was free tyrosine, and not produced by hydrolysis during extraction of tyrosine gallates. These are secondary metabolites that are abundant in some species of *Inga*, and are composed of one or more gallic acids esterified to tyrosine, presumably via galloyltransferases (Fig. S2 and Table S1; Lokvam, Clausen, Grapov, Coley, & Kursar, 2007). To confirm this, two expanding leaf samples of *I. laurina*, a species known to contain tyrosine gallates, were extracted using the tyrosine extraction method described above. This was compared to a 15-minute extraction at room temperature using 60% acetonitrile, a method that would not cause hydrolysis. The four extractions were analyzed using UPLC-MS in negative mode. The ion abundances of the tyrosine gallate peaks were comparable between extraction methods demonstrating that our quantification of free tyrosine was not influenced by the presence of tyrosine gallates (see Supporting Information; Figs. S3 to S6 and Table S2).

We also quantified the amounts of tyrosine gallates and the chemical analogs: tyramine gallates (Davis et al. 2007; Fig. S2, S7, S8 and Tables S1, S2) and DOPA-gallates (Fig. S2, S9) using UPLC-MS analysis in positive mode and then searching the chromatograms for the *m/z* values associated with each molecule (Supporting Information and Tables S3, S4).

Phylogenetic signal

We tested for phylogenetic signal in tyrosine overexpression using Pagel's lambda (λ, Pagel, 1999). Lambda varies from 0 to 1, with values close to 0 indicating no phylogenetic signal (close relatives are not more similar than distant relatives), and values close to 1 suggesting that the trait is evolving according to Brownian motion. For this, we used the principal coordinates of the tyrosine expression distance matrix in the R (version 3.4.3) package *phytools* (Revell, 2018).

Ancestral state reconstruction

We used maximum-likelihood ancestral state reconstruction to determine whether the overexpression of tyrosine and the expression of derived secondary metabolites can be explained by a single origin and to identify the state of the most recent common ancestor for this derived clade. The reconstruction was performed on the preliminary unpublished maximum-likelihood analysis of the hybrid-capture data for the whole genus *Inga.* The ancestral states were reconstructed on a maximum-likelihood framework using the ace command in the R package *phytools* (Revell, 2018). Two transition rate matrices were considered, a two-parameter model that had different rates for forward and reverse transitions (ARD) and a single-parameter model that assumed forward and reverse transition rates were equal (ER). The model fit was examined using the Akaike Information Criteria, with the fitDiscrete command from the package geiger v2.0.6 (Harmon et al., 2016).

Herbivore collections

Each day during the field season we walked trails and scanned ~20m on each side for *Inga* saplings producing young leaves. We visually searched ~30 young-leaf flushes per *Inga*

species and collected only those larvae that were found feeding. Herbivore sampling was performed in all the leaf collection sites, except Brazil. All herbivores were assigned to morphospecies in the field and subsequently to MOTUS (Molecular Operational Taxonomic Units) in the laboratory using sequences from the mitochondrial gene cytochrome oxidase (*COI*) following methods described in Endara et al. (2017). We recorded a total of 5740 individuals in 522 MOTUS from 13 families of insect herbivores. Singleton species were eliminated from all statistical analyses.

We quantified the amount of herbivore damage that occurred during the expansion phase by estimating the percent of leaf area missing on all leaves that had recently flushed on all saplings encountered during our field surveys. Leaves were fully expanded, toughened, but still not fully green, traits indicative of having reached full expansion within the last month. A single observer (Kursar) did visual estimates of percent damage at all sites. On average, 55 individual plants and 4 individual leaves per plant were measured for each species. To compare damage and herbivore abundance for species that contained tyrosine and those that did not, we ran linear mixed models (LMM) in the R package *lme4* (Bates, Maechler, Bolker, & Walker, 2015). Site of collection was included as a random factor in our models. We then performed Type II Wald chi-square ANOVAs on the LMM to test the significance of the models in the R package *car* (Fox & Weisberg, 2011).

Detailed analyses were performed for the three most abundant groups of herbivores. For each of these groups we used maximum-likelihood estimation to model the probability of occurrence using a binomial distribution with the number of trials equal to the total number of herbivore species associated with each *Inga* species. We fitted models that incorporated only the intercept (the simplest model), the effects of tyrosine and galloylated tyrosine compounds, and the principal coordinates of the *Inga* phylogenetic distance matrix. This analysis allowed us to determine the effect of chemical traits (tyrosine and its derivatives), *Inga* host phylogeny and its covariation on the occurrence of insect herbivores. We used the R packages *bbmle* v.1.0.20 (Bolker, 2017) and *emdbook* 347 v.1.3.9 (Bolker, 2016). We performed model comparison based on Akaike´s Information Criterion for small sample sizes (AICc).

Transcriptome

We used transcriptomics to investigate how tyrosine can accumulate to such high levels. It is well established that arogenate-specific dehydrogenase (ADH), the plastidic tyrosine biosynthetic pathway found in most plants, is feedback inhibited by tyrosine (Schenck & Maeda, 2018). In contrast, legumes also have a cytosolic pathway in which the tyrosine biosynthetic enzyme, prephenate dehydrogenase (PDH), lacks strong feedback inhibition by tyrosine (Schenck et al., 2017). Hence, transcriptomics allows investigation of relative changes in the expression of the tyrosine-sensitive and tyrosine-insensitive biosynthetic pathways.

Transcripts of 22 individual plants representing 11 species were sequenced. Expanding leaves between 5-10% of full size were collected, because this stage of expansion is an active period of chemical-defense biosynthesis (Brenes-Arguedas et al, 2006). Leaves were shredded, vacuum-infiltrated using a syringe and preserved in RNAlater (Invitrogen) in the field immediately upon collection. For *I. sapindoides*, total RNA was extracted using Qiagen RNeasy kits with buffer RLC. The other samples were extracted with Invitrogen Plant RNA extraction reagent and the RNA precipitated with 8M LiCl, as in Nicholls et al. (2015). Quality and quantity of RNA were analyzed spectrophotometrically by NanoDrop and on gels and the RNA

was submitted to Edinburgh Genomics, Edinburgh, UK. Libraries were made using Illumina's TruSeq RNA Sample Prep kits.

Accessions from *I. umbellifera* (umb_BCI_180, umb_BCI_189, umb_BCI_013) were sequenced on an Illumina miSeq (250bp reads, 2672496 - 9832649 reads per sample). All other samples were sequenced on an Illumina HiSeq (75bp reads, 23428610 - 35637400 reads per sample) (Table S5). Reads were quality trimmed using Trimmomatic v0.30 (Bolger, Lohse, & Usadel, 2014) with settings: ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36. Trinity (release r20120608, Grabherr et al., 2011) was used for individual de-novo assemblies. Transcriptome annotation used Trinnotate v3.1.0 (https://trinotate.github.io/) in Cyverse (http://www.cyverse.org/atmosphere). Salmon was used to produce transcripts per million reads (TPM) values for each contig (Patro, Duggal, Love, Irizarry & Kingsford, 2017).

Orthologs of the tyrosine-sensitive enzyme (ADH), and the tyrosine-insensitive enzyme (PDH), were identified by blast match to characterized legume orthologs from *Medicago* and *Glycine*. Alignments between reference sequences from *Medicago truncatula* and *Glycine max* and the *Inga* candidates were made in MAFFT 7.4 (Katoh & Stanely, 2013) to confirm orthology. We confirmed the presence of the key residues for tyrosine insensitivity.

Results

Overexpression of tyrosine in expanding leaves evolved once

We quantified tyrosine levels in expanding leaves of 97 *Inga* species spread across the phylogenetic tree and identified a single clade that overexpressed tyrosine (Fig. 1). This clade includes 17 species that overexpress tyrosine and four species that do not, representing three independent losses. Tyrosine overexpression also showed a high phylogenetic signal (PCoA axis 1: $\lambda = 0.9$, $p<0.0001$; PCoA axis 2: $\lambda= 0.92$, $p < 0.0001$). For the species that overexpress tyrosine, investment in this single compound is substantial, averaging 11% DW and ranging from 5% to a maximum investment of 20% in *I. loubryana* (Fig. 1). Model fitting in maximum-likelihood reconstruction supported the ARD model. This model of reconstruction of ancestral states suggests that the most common recent ancestor for this derived clade expressed tyrosine (100% probability) and had a single origin. This places the origin of tyrosine overexpression at 4.47mya (95% credible intervals of 4.13-4.87; Koenen unpubl.) and includes the species that accumulate tyrosine and the sister *capitata* clade (Fig. 1 arrow).

For 11 of the species that overexpress tyrosine, we examined how tyrosine concentrations changed during leaf development (Fig. S12). In all cases, overexpression of tyrosine is only found in expanding leaves, while concentrations in mature leaves of the same species were almost always below the detection limit of the assay (<0.36% DW). In fact, only two of the 44 mature-leaf samples had measurable levels of tyrosine (one each for *I. alba* (3.9%) and *I. heterophylla* (0.8%)*;* Table S4*).* Other individuals from these two species did not have measurable levels of tyrosine. More sensitive assays show that levels in mature leaves are probably <0.05% DW (Lokvam et al., 2006).

Overexpression of tyrosine can lead to synthesis of secondary metabolites

Six *Inga* species have abundant tyrosine gallates, phenolic compounds in which one or more gallic acid moieties are esterified to tyrosine (Lokvam et al., 2007). These include mono-, di-, tri- and tetra-gallates (Fig. S2, Table S4). All of these species overexpress tyrosine, but

represent various lineages within the tyrosine clade. Of the 97 sampled *Inga* species, none outside the tyrosine clade produce tyrosine gallates.

We also found small amounts of DOPA mono- and di-gallates in *I. obidensis, I. laurina,* and *I. heterophylla* (Fig. S2, Table S4). DOPA is biosynthesized from tyrosine in a single step by the 3-hydroxylation of the phenolic ring by tyrosine 3-monooxygenase. The DOPA gallates are analogous to tyrosine gallates, with gallic acid moieties esterified to DOPA. We have not detected significant amounts of DOPA gallates in any other *Inga* species. Even though *I. obidensis*, *I. heterophylla* and *I. laurina* are very similar in chemistry (all have abundant free tyrosine, tyrosine gallates and low levels of DOPA gallate; Fig. 1), the phylogeny supports the hypothesis that these are not sister taxa and belong to three discrete clades. Hence, it appears that their shared gallate chemistry may have evolved independently.

Interestingly, five of the species in the *I. capitata* complex produce tyramine gallates but no free tyrosine, even though the most recent common ancestor of the *I. capitata* and tyrosine clade expressed tyrosine (Fig. 1). Tyramine is a primary metabolite that is biosynthesized in a single step by decarboxylation of tyrosine and the gallates are analogous to the tyrosine and DOPA gallates (Davis et al. 2008; Fig. S2). Although six of these species were originally considered a single species, *I. capitata* (Pennington 1997), they differ based on both chemistry and DNA sequences (Endara et al. 2015). The elevated levels of tyramine gallates, but not of free tyrosine nor its gallates, suggest that regulation of biosynthesis operates differently in these species than for the tyrosine clade.

None of the *Inga* species outside the tyrosine and *capitata* clades (Fig. 1) contained derivatives of tyrosine, such as tyramine, DOPA or the gallates. This is consistent with the expectation that overexpression of tyrosine is a first step in the evolution of more derived secondary metabolites.

Higher transcripts of tyrosine-insensitive biosynthetic enzymes in Inga that overexpress tyrosine

The accumulation of high concentrations of free tyrosine presents questions regarding the regulation of metabolite expression. In most plants, negative feedback by tyrosine to a key biosynthetic enzyme, arogenate-specific dehydrogenase (ADH), prevents over-accumulation of tyrosine (Schenck & Maeda 2018). But, many legumes have a second biosynthetic pathway, prephenate-specific dehydrogenase (PDH), that is insensitive to tyrosine (Schenck et al., 2017). Although ADH and PDH are very closely related, their major difference is a single amino acid that determines both substrate specificity and the inhibition of enzyme activity by tyrosine. Hence, we examined the transcriptome of expanding leaves of three species that overexpress tyrosine and five distantly related species that lack overexpression, plus a single species from the tyrosine clade that shows a secondary loss of tyrosine overexpression, "*cf. umbellifera*" (Fig. 2). We found that all species of *Inga* express ADH, as recognized by a crucial aspartic acid codon at amino acid residue 213. As expected, we found that ADH shows low levels of transcription across *Inga* species regardless of tyrosine expression level. In fact, the three tyrosine-producing species have some of the lowest levels of ADH transcripts. However, transcripts of the tyrosine-insensitive PDH have intermediate to high abundance in the species that overexpress tyrosine compared to species that do not. This contrast is even found in two closely related species in the *I. umbellifera* complex, *cf. umbellifera* from French Guiana has no detectable levels of tyrosine and higher transcript levels of the more sensitive enzyme (ADH) along with low expression of PDH, while *I. umbellifera* from Panama has 11% tyrosine and high levels of the tyrosine-insensitive enzyme (PDH).

Another species for which we have transcripts, "*capitata*12-FG" from French Guiana, also shows high levels of the insensitive PDH enzyme (Fig. 2). This species is in the *I. capitata* clade, which is sister to the tyrosine clade (Fig. 1). Although "*capitata*12" does not overexpress tyrosine, it makes tyramine gallates, a secondary metabolite downstream from tyrosine. Another species in the clade, "*capitata*33" contains low concentrations of tyramine gallates (Table S4), and has intermediate levels of the PDH transcript (Fig. 2). This suggests that those species in the *I. capitata* clade with high levels of tyramine derivatives may depend upon the tyrosine-insensitive pathway for the biosynthesis of tyramine.

Overexpression is more likely in species that suffer high herbivory

We ask why overexpression has evolved, given that it is costly in terms of nitrogen investment, and although experiments suggest it is toxic, it is likely not as effective against herbivores as more derived secondary metabolites (Lokvam et al., 2006). To test our hypothesis that overexpression would only be favored in species that have less effective secondary metabolites, we quantified levels of damage in the field to expanding leaves. High levels of herbivory indicate that the secondary metabolites are not effective against the suite of herbivores that can potentially attack them. As predicted, species that overexpress tyrosine accrued significantly higher levels of damage during leaf expansion than species without tyrosine ($p<0.001$, Fig. 3). This was true at all four sites, even though sites differed in the mean levels of damage, with herbivory being highest in Peru. We did not find a significant effect of tyrosine overexpression on the number of individual herbivores found on each plant. In the discussion, we outline why higher herbivory is most likely due to less effective secondary metabolites rather than attraction to tyrosine.

It is not possible to robustly test if the derivatives are more effective in reducing herbivory than tyrosine alone, as few *Inga* make tyrosine derivatives (Fig. 1). The rarity of derivatives could indicate that they are not measurably more effective than free tyrosine. Alternatively, we suggest that there has not been sufficient time for derivatives to emerge, given the recent innovation of tyrosine overexpression (4.47 mya).

Overexpression of tyrosine shapes host choice in insect herbivores

The herbivores attacking *Inga* are predominantly Lepidoptera, with most species belonging to three groups: the families Lycaenidae and Noctuidae, and the superfamily Gelechioidea. Detailed analyses for these families of herbivores showed that these groups have diversified responses to the overexpression of tyrosine and its derived secondary metabolites in their *Inga* hosts. For Lycaenidae, their probability of occurring on a host increased significantly if the host overexpressed tyrosine (Table 1, Fig. S13). For example, the proportional odds value of 1.76 indicates that Lycaenidae are 76% more likely to occur on a host if it contained free tyrosine. In contrast, the presence of tyrosine and tyramine gallates negatively affected their occurrence, although not significantly (proportional odds <1.0, Table 1). Noctuidae responded differently. Their presence was negatively associated with free tyrosine, but they were more abundant on *Inga* species that express gallates of both tyrosine and tyramine (Table 1, Fig. S14). For the superfamily Gelechioidea, the presence of tyrosine or any tyrosine-derived compound marginally reduced their probability of occurrence (Table 1, Fig. S15).

Discussion

Single origin of tyrosine overexpression

Our analysis of the whole genus *Inga* points to a single origin of overexpression for tyrosine with several secondary losses. Reconstruction of ancestral states suggests that this biosynthetic innovation occurred at the base of the tyrosine and *capitata* clades, approximately 4.5 mya (Fig. 1). This very likely involved the upregulation of the tyrosine-insensitive enzyme, PDH. The fact that free tyrosine does not accumulate in the *I. capitata* complex indicates that, for this clade, regulation of biosynthesis has favored the accumulation of tyrosine derivatives over tyrosine itself.

The strong phylogenetic signal for tyrosine overexpression differs from the patterns seen for *Inga* secondary metabolites (Kursar et al., 2009, Endara et al., 2017, Endara et al., 2018). Previously we have argued that herbivores have selected for divergent defenses, such that close relatives are not more similar in secondary metabolites (Coley et al., 2018). The divergence among *Inga* species arises primarily through different combinations of a shared arsenal of phenolics and saponins. In other words, all *Inga* likely have the structural genes to make most of the compounds, but they are up- or down-regulated in different combinations in different species. This mode of generating chemical diversity contrasts with key but rare innovations arising from changes in the genes that code for biosynthetic enzymes giving rise to novel molecules, such as with glucosinolates or latex (Berenbaum, 1981; Ryan et al., 2012). Overexpression of tyrosine in *Inga* is likely not due to novel enzymes, as the tyrosine-insensitive enzyme, PDH, was detected in all *Inga* surveyed (Fig. 2) and is present in most members of the legume family (Schenck & Maeda 2018). Instead, differential *regulation* of key enzymes is the most likely explanation, a pattern similar to that for the secondary metabolites (Coley et al., 2018). Nonetheless, a change in the accumulation patterns of tyrosine only emerged once in *Inga* suggesting that it is a relatively rare event.

We have emphasized phylogenetic conservatism in that tyrosine/tyramine overproduction evolved only once. Nevertheless, a number of lineages within the tyrosine/tyramine clade produce neither free tyrosine nor tyrosine/tyramine gallates (Fig. 1), indicating some evolutionary lability in the loss of overproduction. Furthermore, within this clade, different species have divergent phenolic and saponin chemistry. Thus, although tyrosine overexpression is phylogenetically conserved, the expression of other secondary metabolites varies considerably among members of this tyrosine-accumulating clade and across *Inga* in general.

Tyrosine is overexpressed primarily in expanding leaves

Expanding leaves of most tropical species are high in protein due to the fact that they are growing, and low in fiber as secondary cell walls cannot be laid down until the cells have stopped expanding (Kursar & Coley, 2003). As a consequence, expanding leaves are tender and nutritious, and preferred by herbivores. Once leaves are fully expanded and can be physically defended by toughness, the chemical defenses so critical for the expansion phase are not needed to the same extent. Thus, at the end of leaf expansion, it would be advantageous to 'recycle' unneeded chemicals. This may not be possible for some compounds, such as condensed tannins. Tyrosine, on the other hand, is readily catabolized or diverted to other functions, with levels dropping in just a short time after full expansion (Lokvam et al., 2006; Bixenmann, Coley, Weinhold, & Kursar 2016). Recycling tyrosine or other primary metabolites could reduce the costs of investing limiting nitrogen in defense. Although hyper-accumulation of tyrosine may be physiologically possible in mature leaves, it is apparently not adaptive.

Overexpression of tyrosine and herbivore host use

Our results are consistent with the idea that the overexpression of tyrosine exerts constraints on herbivore host use. Tyrosine and galloylated tyrosine-derived compounds were consistently selected as important predictors for host associations across the three most abundant insect families (Table 1). In fact, the *Inga* species that have lost the capacity to express tyrosine or any of its derived secondary metabolites (*I. laterifolia, I. paraensis and I. stipularis*, are attacked by generalist herbivores that feed on other *Inga* hosts that also do not express tyrosine or any of its derivatives. Furthermore, our analyses also suggest that the response to tyrosine overexpression varies greatly among herbivore species that have evolved different life history traits. Much of this variation is undoubtedly related to the fact that specialized insect herbivores can adapt over evolutionary time to the defenses expressed by their hosts (Agrawal & Fishbein 2008). For example, larvae from Lycaenidae are less likely to be associated with *Inga* that contain tyrosine- or tyramine-gallates, and were the only family that showed preference for hosts that overexpress tyrosine (Fig. S13). Lycaenidae represents a special case because many species of this family have evolved adaptations to hijack the plant's defenses against herbivores for their own benefit. Besides causing copious amount of damage to young tissue, Lycaenidae larvae use the plant's ant-guards for their own protection. This conflict might be even stronger because they also feed on the extrafloral nectar that *Inga* uses to attract ants (DeVries & Baker 1989). They are known to be attracted to plants with high levels of amino acids in order to meet their own energetic requirements and to provide rewards for the attendant ants (Pierce 1985, Pellisier et. al., 2012). Their positive association with tyrosine-producers is therefore consistent with this suite of unique adaptations.

In contrast, Noctuidae tend to avoid hosts that contain tyrosine, but are overrepresented on *Inga* with tyrosine-and tyramine-gallates (Fig. S14). Our previous studies on *Inga* in Peru showed that the presence of tyrosine derivatives was among the most important factors for host selection by Noctuidae (Endara et al., 2017). In the present study, with a larger dataset we show that this pattern extends to their entire regional distribution (Panama, Guiana Shield and Amazonia). For Gelechioidea, tyrosine expression is marginally significant, with Gelechioidea larvae avoiding *Inga* hosts that produce either this compound or its derivatives (Fig. S15). Thus, for lepidopteran herbivores associated with *Inga*, host choice might be constrained at the family level, with each group having different sets of adaptations for handling tyrosine or its derivatives.

Other evidence also suggests that the overexpression of tyrosine and tyrosine-related compounds has shaped the evolution of herbivore host use and diversification. For leaf-feeding sawflies of the family Argidae, another group that causes substantial herbivory on *Inga* and that is highly specialized (each species feeding on one or two *Inga* hosts), the evolution of host range and species divergence has been driven by the production of galloylated tyrosine-derived compounds in their *Inga* hosts (Endara et al., 2018). Thus, it is very likely that the overexpression of tyrosine has played an important role in the processes shaping host association and species divergence both at local and regional scales.

Is tyrosine overexpression a defense?

Although the effects of overexpression of primary metabolites on herbivores are poorly studied, we propose several lines of argument suggesting that the high concentrations of L-tyrosine seen in *Inga* (5-20%) likely are functioning as a defense against the majority of herbivores. First, as outlined in the introduction, feeding trials with insects and mammals show toxic or lethal effects (Rehr et al., 1973, Hazel et al., 2003; Lokvam et al, 2006, De Prá et al. 2014, Sterkel et al. 2016). Second, if tyrosine was generally nutritive/attractive to herbivores, this

mutation would quickly have been eliminated. Third, our data show that most herbivores have negative associations with tyrosine-producers. The only family showing a positive association with tyrosine-producers (Lycaenidae), appears to have evolved mechanisms for circumventing this defense and exploiting the system.

The fact that herbivores attack tyrosine accumulators more, could mean that they have a less effective secondary metabolite profile, or that tyrosine is an attractant. Although we cannot rule out the latter, based on the literature, our experiments, and our findings in the field, we think that the most parsimonious explanation for tyrosine overexpression is that it is generally toxic or lethal to non-adapted herbivores.

Overexpression of primary metabolites

It is unusual to see overexpression of primary metabolites, even those that may function as a defensive metabolite. To our knowledge, the only other plant to overexpress tyrosine is a species of bamboo, with young shoots containing 30% DW tyrosine (Nomura & Yamada, 1974). We propose four non-exclusive hypotheses for why the accumulation of high levels of primary metabolites is rare.

First, it is possible that many primary metabolites would not be toxic, and overexpression would quickly be selected against, particularly if they were less toxic than existing secondary metabolites. In a survey of 20 amino acids incorporated into artificial diets at 1% DW, only eight reduced the probability of survival to adulthood in larvae of a bruchid beetle (Janzen, Juster & Bell, 1977). When concentrations were increased to 5% DW, an additional seven amino acids produced significant effects. L-tyrosine belonged to this latter group. So, based on toxicity to herbivores, there are at least several amino acid candidates for overexpression. However, in order to accumulate high concentrations of amino acids, mutations to the biosynthetic enzymes must occur to reduce negative feedback. In the case of *Inga*, this is accomplished by the tyrosine insensitive enzyme, PDH.

A second possible explanation for the rarity of overexpression of primary metabolites may be the cost in terms of nitrogen. Allocating limited nitrogen to defense rather than growth may not be adaptive. The levels of tyrosine in the young leaves of *Inga* ranged from 5-20% of the leaf DW. As tyrosine contains nitrogen, an investment of 10% tyrosine is equivalent to 0.8% nitrogen. Given that a typical expanding leaf without tyrosine has about 3% of the leaf DW as nitrogen (Kursar & Coley, 1991), this represents an additional investment of 27% in nitrogen - a considerable portion of the total nitrogen budget. Herbivores remove ~30% of the leaf area in the few weeks that leaves are expanding, so species without tyrosine would lose 0.9% DW of their nitrogen investment. For tyrosine accumulators, which have 13% DW nitrogen (3% plus ~10%), they will lose four times that amount to herbivores (3.9%).

Although *Inga* has symbiotic relationships with nitrogen-fixing *Rhizobia*, we doubt this can provide sufficient nitrogen to compensate for the extra investment. *Inga* saplings growing in the shaded understory have very low levels of nodule formation (Kursar & Coley pers. obs.), perhaps because the carbon cost of maintaining *Rhizobia* is high (Hedin, Brookshire, Menge, & Barron, 2009). Barron, Purves & Hedin (2011) also found nodulation to be near zero in canopy trees of 11 *Inga* species in an old growth lowland tropical forest, including one that accumulates tyrosine. However, it would be interesting to more broadly test if tyrosine accumulators have evolved a closer association with *Rhizobia* than non-accumulators. For species that do not have symbioses with *Rhizobia*, the opportunity costs of investing in tyrosine would be even higher.

A third hypothesis is that conditions in which overexpression is beneficial are rare. For

selection to favor overexpression it must be at least as cost-effective against herbivores as a similar investment in the secondary metabolites produced by that species. In the *Inga* system, tyrosine investments are high (5-20% DW) and although tyrosine did reduce larval growth, it is not particularly effective compared to other compounds (Lokvam et al., 2006). For example, nicotine, a nitrogen containing secondary metabolite is present and effective at levels 10-20 times less than tyrosine (Baldwin 1999). Thus, conditions where the cost/benefit value compares favorably to that for secondary metabolites would only occur if the secondary metabolites were no longer very effective. Our data showing higher herbivory on species with tyrosine overexpression is consistent with this prediction. Given that overexpression only evolved once, it appears that favorable conditions might not be that common.

A fourth possibility is that selection would favor the rapid evolution of derived secondary metabolites, such that we rarely catch instances of overexpression of primary metabolites. Indeed, tyrosine is the initial compound in numerous biosynthetic pathways leading to a diversity of secondary metabolites (Schenck & Maeda, 2018), including the galloylated compounds found in *Inga*. Presumably, these more biologically derived compounds would be more effective than primary metabolites. We therefore predict that selection would favor derived compounds that are both less costly and more effective as defenses (McKey 1974, Rhoades 1979). The fact that only a few *Inga* species have taken the next evolutionary step of making tyrosine-derived secondary metabolites, and that all of these species are in the tyrosine clade, may be due to the fact that we are capturing a relatively recent innovation (4.47 mya). We presume that over a longer evolutionary trajectory, tyrosine would be replaced by derived secondary metabolites with greater efficacy against herbivores. Thus, the first evolutionary step in the production of novel secondary metabolites in *Inga* is likely upregulation of tyrosine, perhaps through gene duplication. This may be a common first step in many species, as no novel enzymes are needed to up-regulate an existing metabolic pathway, and the required mutations are less likely to be deleterious than mutations to coding regions. Once duplicated, subsequent modification of the primary metabolite potentially allows independent evolution of defensive compounds.

Acknowledgements

This work could not have been done without valuable field assistance provided by Marjory Weber, Emily Kearney, Wilmer Rosendo, Carine Emer, Georgia Sinimbu, Wilder Hidalgo, Julio Grandez, Joe Sixto Saldaña, Zachary Benavidez, Allison Thompson, Yamara Serrano and Mayra Ninazunta. We thank John Lokvam for prelimnary chemical analyses. The research was supported by grants from the National Science Foundation (DEB-0640630 and DIMENSIONS of Biodiversity DEB-1135733) and the Nouragues Travel Grants Program, CNRS, France to PDC and TAK and UK BBSRC/NERC SynTax to RTP, PDC and TAK. This work was greatly facilitated by research permits from the governments of Panama, Peru, Brazil and Ecuador.

Author contributions statement

PDC, TAK, and M-JE designed the project and conducted field research. GG, AGM, AJS and TAK did the chemical analyses and CAK did the transcriptomics. JAN, RTP and CAK contributed the next-generation DNA sequence data and phylogenies. PDC, TAK and M-JE wrote the manuscript.

Data accessibility

Data available from the Dryad Digital Repository https://doi.org/xxxx (data will be deposited at Dryad once the manuscript is accepted).

Tables

Table 1. Results of the best-fit maximum likelihood model for the three most abundant lepidopteran families against tyrosine and galloylated tyrosine-derived compounds. Proportional odds estimate is the proportional change in odds for the probability of occurrence for each family per unit of its predictor variable. A value of 1.0 indicates no effect, while values >1.0 indicate the proportional increase of larval occurrence in response to the predictor variable, and values <1.0 indicate the proportional decrease of larval occurrence. Significance codes: *ns*= not significant, *≤0.06, **≤0.05, ***≤0.01, ****≤0.001.

Herbivore	Predictor Variable	Proportional odds estimate	95% Confidence Interval	*P-value*
Lycaenidae	Tyrosine	1.76	2.76-1.06	***
	Tyrosine gallates	0.68	1.38-0.3	ns
	Tyramine gallates	0.49	1.06-0.18	ns
Noctuidae	Tyrosine	0.76	1.37-0.39	ns
	Tyrosine gallates	1.86	3.26-1	**
	Tyramine gallates	2.68	4.49-1.51	****
Gelechioidea	Tyrosine and derived compounds	0.76	1.05-0.53	*

Figures

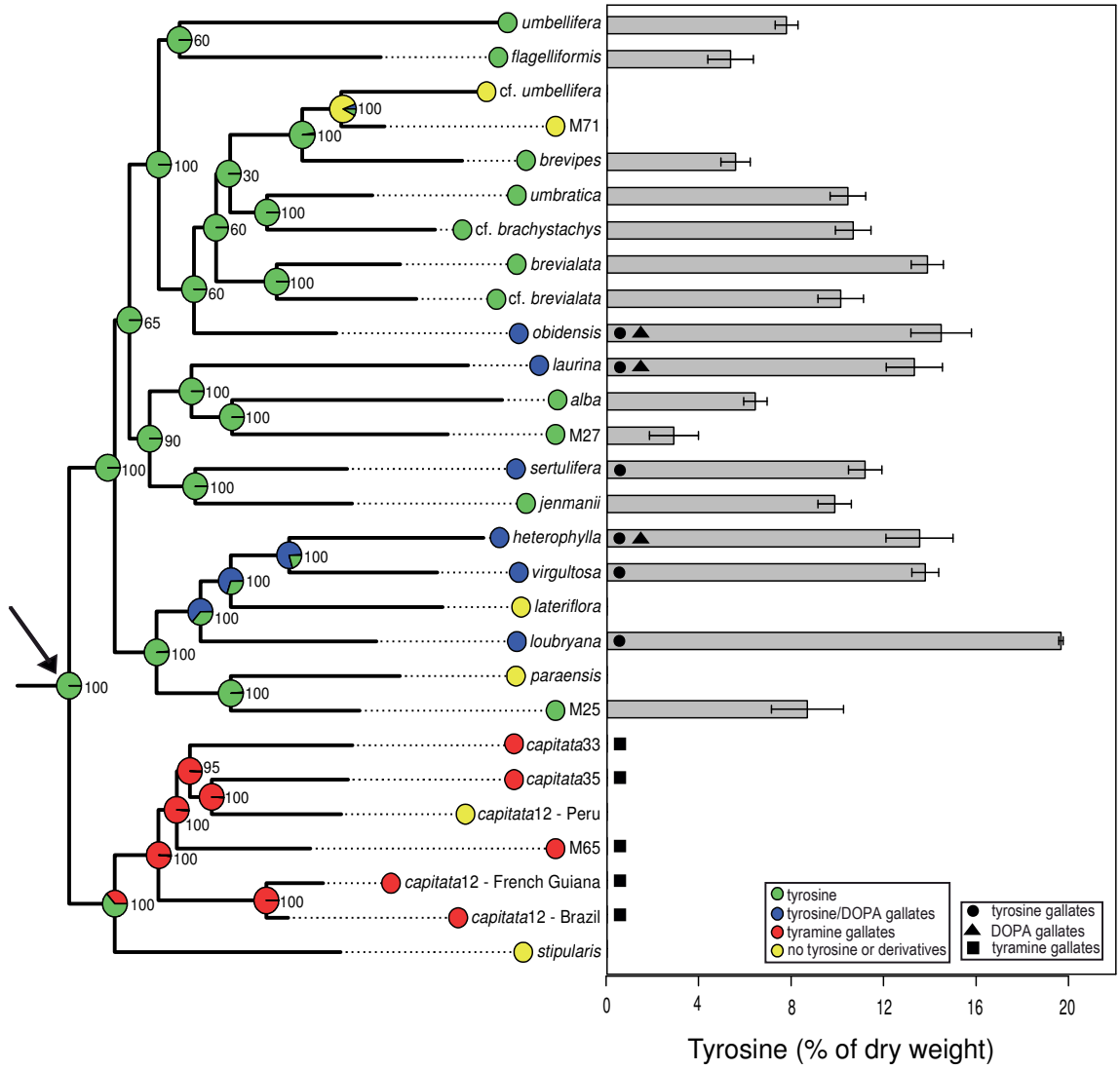

Fig. 1. Maximum-likelihood estimates of phylogenetic relationships within the *Inga* clade that includes all the species that overexpress tyrosine and derivatives, as well as the percent dry weight of tyrosine in the expanding leaves. Symbols indicate the presence of moderate or high levels of tyrosine gallates (circles) and tyramine gallates (squares) and low levels DOPA gallates (triangles). Species with very low levels of gallate expression were not included, but values are presented in Table S4 for each species and site. The colored pie diagrams indicate proportions of character histories that reconstructed tyrosine (green), tyrosine derivatives (blue and red), or non-tyrosine related chemistry (yellow) as the ancestral state for the most recent common ancestor of all the tyrosine species clade. The arrow indicates the likely origin of the overexpression of tyrosine at 4.47 mya.

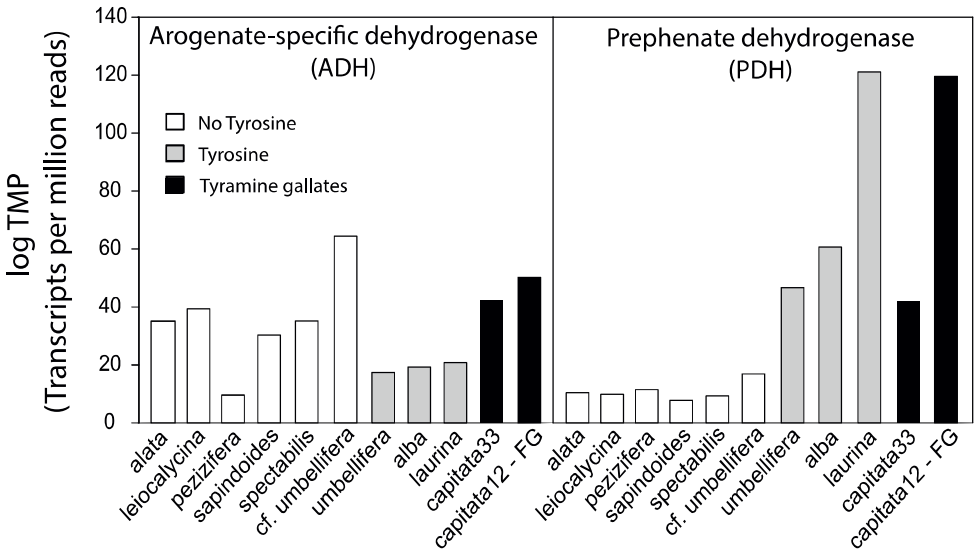

Fig. 2. Estimates of transcript abundance for the tyrosine-sensitive enzyme, arogenate-specific dehydrogenase (ADH) and the tyrosine-insensitive enzyme, prephenate dehydrogenase (PDH) for expanding leaves of *Inga* species that do not overexpress tyrosine (open bars), that do overexpress tyrosine (grey bars) and that make tyramine gallates (black bars).

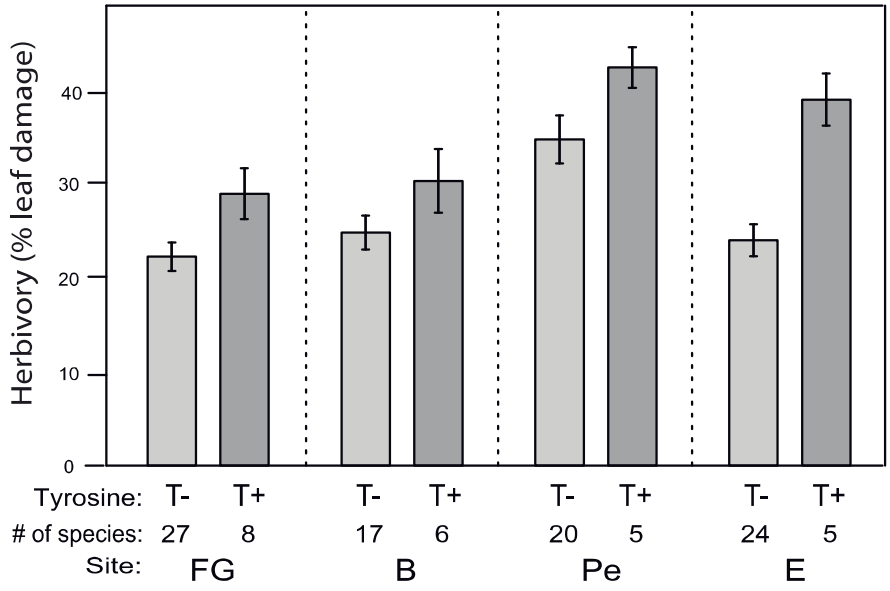

Fig. 3. Herbivore damage to expanding leaves for species with and without tyrosine overexpression. Species that contained tyrosine-derivatives in addition to free tyrosine were included in the tyrosine-plus category. Data are from French Guiana (FG), Brazil (B), Peru (Pe) and Ecuador (E). Panama is excluded from the analysis as only one species contained tyrosine. An ANOVA exploring the effect of tyrosine status (T+ or T-) on herbivory showed lower damage for species without tyrosine ($p < 0.001$)

References

Agrawal, A.A. & Fishbein, M. (2006) Plant defense syndromes. Ecology, 87, S132–S149. doi: 10.1890/0012-9658(2006)87[132:PDS]2.0.CO;2

Baldwin, I.T. (1999) Inducible nicotine production in native *Nicotiana* as an example of adaptive phenotypic plasticity. Journal of Chemical Ecology, 25, 3-30. doi: 10.1023/A:1020880931488

Barron, A.R., Purves, D.W., & Hedin, L.O. (2011) Facultative nitrogen fixation by canopy legumes in a lowland tropical forest. Oecologia, 165, 511-520. doi: 10.1007/s00442-010-1838-3

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67, 1-48. doi: 10.18637/jss.v067.i01

Becerra J.X. (1997) Insects on plants: Macroevolutionary chemical trends in host use. Science, 27, 253-256. doi: 10.1126/science.276.5310.253

Berenbaum, M.R., & Feeny, P. (1981) Toxicity of angular furanocoumarins to swallowtail butterflies: Escalation in a coevolutionary arms race? Science, 212, 927-929. doi: 10.1126/science.212.4497.927

Bixenmann, R.J., Coley, P.D., Weinhold, A., & Kursar, T.A. (2016) High herbivore pressure favors constitutive over induced defense. Ecology and Evolution, 6, 6037-6049. doi:10.1002/ece3.2208

Bolger, A. M., Lohse, M., & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Bolker, B. (2016) emdbook: Ecological models and data (book support). Available at: http://cran.r-project.org/web/packages/emdbook/emdbook.pdf

Bolker, B. (2017) bbmle: Tools for general maximum likelihood estimation. Available at: http://cran.r-project.org/web/packages/bbmle/bbmle.pdf

Brenes-Arguedas, T., Horton, M.W., Coley, P.D., Lokvam, J., Waddel, R.A., Meizoso- O'Meara, B.E. & Kursar, T.A. (2006) Contrasting mechanisms of secondary metabolite accumulation during leaf development in two tropical tree species with different leaf expansion strategies. Oecologia, 149, 91-100. doi.org/10.1007/s00442-006-0423-2

Coley P.D., & Kursar T.A. (2014) On tropical forests and their pests. Science, 343, 35-36. doi: 10.1126/science.1248110

Coley, P.D., Endara, M-J., & Kursar, T.A. (2018) Consequences of interspecific variation in defenses and herbivore host choice for the ecology and evolution of *Inga*, a speciose rainforest tree. Oecologia, 187, 361-376. doi: 10.1007/s00442-018-4080-z

Davis, R.A., Simpson, M.M., Nugent, R.B., Carroll, A.R., Avery, V.M., Rali, T., … Quinn, R.J., (2007) Pim2 inhibitors from the Papua New Guinean plant *Cupaniopsis macropetala*. Journal of Natural Products, *71*, 451-452. doi: 10.102/np070431w

DeVries, P.J. & Baker, I. (1989) Butterfly exploitation of an ant-plant mutualism: Adding insult to herbivory. Journal of the New York Entomological Society, 97, 332-340.

De Prá, S.D.T., Ferreira, G.K. Carvalho-Silva, M., Vieira, J.S., Scaini, G., Leffa, D.D, … Streck, E.L. (2014) L-Tyrosyine induces DNA damage in brain and blood of rates. Neurochemical Research, 39, 202-207. doi: 10.1007/211064-013-1207-9

Dexter, K.G., Lavin, M., Torke, B.M., Twyford, A.D., Kursar, T.A., Coley, P.D., … Pennington, R.T. (2017) Dispersal assembly of rain forest tree communities across the Amazon basin. Proceedings of the National Academy of Sciences (USA), 114, 2645-2650. doi:10.1073/pnas.1613655114

Endara, M-J., Weinhold, A., Cox, J.E., Wiggins, N.L., Coley, P.D., & Kursar, T.A. (2015) Divergent evolution in antiherbivore defences within species complexes at a single Amazonian site. Journal of Ecology, 103, 1107-1118. doi:10.1111/1365-2745.12431

Endara, M-J., Coley, P.D., Ghabash, G., Nicholls, J.A., Dexter, K.G., Donoso, D.A., … Kursar, T.A. (2017) Coevolutionary arms race versus host defense chase in a tropical herbivore–plant system. Proceedings of the National Academy of Sciences (USA), 114, E7499-E7505. doi:10.1073/pnas.1707727114

Endara, M-J., Nicholls, J.A., Coley, P.D., Forrister, D.L., Younkin, G.C., Dexter, K.G., … Kursar, T.A. (2018) Tracking of host defenses and phylogeny during the radiation of neotropical *Inga*-feeding sawflies (Hymenoptera; Argidae). Frontiers in Plant Science 9:1237. doi: 10.3389/fpls.2018.01237

Ehrlich, P.R., & Raven, P.H. (1964) Butterflies and plants: A study in coevolution. Evolution, 18, 586-608. doi:10.2307/2406212

Fine, P.V.A., Metz, M.R., Lokvam, J., Mesones, I., Zuñiga, J., Lamarre, G., … Baraloto, C. (2013a) Insect herbivores, chemical innovation, and the evolution of habitat specialization in Amazonian trees. Ecology 94, 1764-1775. doi:10.1890/12-1920.1

Fine, P.V.A., Zapata, F., Daly, D.C., Mesones, I., Misiewicz, T.M., Cooper, H.F., … Barbosa, C. (2013b) The importance of environmental heterogeneity and spatial distance in generating phylogeographic structure in edaphic specialist and generalist tree species of *Protium* (Burseraceae) across the Amazon Basin. Journal of Biogeography 40, 646-661. doi:10.1111/j.1365-2699.2011.02645.x

Fox, J., & Weisberg, S. (2011). An {R} Companion to Applied Regression, (2nd ed.). Thousand Oaks, CA: Sage.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., …Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology, 29, 644–652. doi: 10.1038/nbt.1883

Harmon, L., Weir, J., Brock, C., Glor R., Challenger, W., Hunt, G., … & Eastman, J. (2018) R package Geiger: Analysis of evolutionary diversification. Accessed on 12/03/2018.

Hazel, M.H., Ianowski, J.P., Christensen, R.J., Maddrell, S.H.P. & O'Donnell, M.J. (2003) Amino acids modulate ion transport and fluid secretion by insect Malpighian tubules. Journal of Experimental Biology 206: 79–91. doi: 10.1242/jeb.00058

Hedin, L.O., Brookshire, E.N.J., Menge, D.N.L., & Barron, A.R. (2009). The nitrogen paradox in tropical forest ecosystems. Annual Review of Ecology, Evolution and Systematics, 40, 613–635. doi:10.1146/annurev.ecolsys.37.091305.110246

Janz, N. (2011). Ehrlich and Raven revisited: mechanisms underlying codiversification of plants and enemies. Annual Review of Ecology, Evolution and Systematics, 42, 71-89. doi:10.1146/ annurev.ecolsys-102710-145024

Janzen, D. H., Juster, H. B., & Bell, E.A. (1977) Toxicity of secondary compounds to the seed-eating larvae of the bruchid beetle *Callosobruchus maculatus*. Phytochemistry, 16, 223–227. doi: 10.1016/S0031-9422(00)86790-4

Katoh, K., & Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolutions, 30, 772-780. doi.org/10.1093/molbev/mst010.

Kraft, N., Valencia, R., & Ackerly, D. (2008). Functional traits and niche-based tree community assembly in an Amazonian forest. Science 332, 580–582. doi: 10.1126/science.1160662

Kursar, T.A., & Coley, P.D. (1991) Nitrogen content and expansion rate of young leaves of rainforest species: Implications for herbivory. Biotropica, 23:141-150. doi:10.2307/2388299

Kursar, T.A., & Coley, P.D. (2003) Convergence in defense syndromes of young leaves in tropical rainforests. Biochemical Systematics and Ecology, 21, 929-949. doi: 10.1016/S0305-1978(03)00087-5

Kursar, T.A., Dexter, K.G., Lokvam, J., Pennington, R.T., Richardson, J.E., Weber, M.G., … Coley, P.D. (2009) The evolution of anti-herbivore defenses and their contribution to species coexistence in the tropical tree genus *Inga*. Proceedings of the National Academy of Sciences (USA), 106, 18073-18078. doi:10.1073/pnas.0904786106

Lokvam, J., Brenes-Arguedes, T., Lee, J.S., Coley, P.D., & Kursar, T.A. (2006). Allelochemic function for a primary metabolite: The case of L-tyrosine hyper-production in *Inga umbellifera* (Fabaceae). American Journal of Botany, 93, 1109-1115. doi:10.3732/ajb.93.8.1109

Lokvam, J., Clausen, T.P., Grapov, D., Coley, P.D., & Kursar, T.A. (2007) Galloyl depsides of tyrosine from young leaves of *Inga laurina*. Journal of Natural Products, 70, 134-136. doi: 10.1021/np060491m

Marquis, R.J., Salazar, D., Baer, C., Reinhardt, J., Priest, G., & Barnett, K. (2016) Ode to Ehrlich and Raven or how herbivorous insects might drive plant speciation. Ecology, 97, 2939-2951. doi:10.1002/ecy.1534

McKey D. (1974) Adaptive patterns in alkaloid physiology. American Naturalist, 108, 305-320. doi.org/10.1086/282909

Nicholls, J.A., Pennington, R.T., Koenen, E. J. M., Hughes, C.E., Hearn, J., … Kidner, C. A. (2015). Using targeted enrichment of nuclear genes to increase phylogenetic resolution in the neotropical rain forest genus *Inga* (Leguminosae: Mimosoideae). Frontiers in Plant Science, 6, 710. doi:10.3389/fpls.2015.00710

Nomura, T., & Yamada, T. (1974) X-ray analysis of tyrosine in growing stage of bamboo (*Phyllostachys edulis* A., & C. Riviere). Wood Research Bulletin of the Wood Research Institute Kyoto University, 56, 21-27.

Pagel, M. (1999) Inferring the historical patterns of biological evolution. Nature, 401, 877–888.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., & Kingsford, C. (2017) Salmon: fast and bias-aware quantification of transcript expression using dual-phase inference Nature Methods, 14, 417–419. doi:10.1038/nmeth.4197

Pellissier, L., Rasmann, S., Litsios, G., Fiedler, K., Dubuis, A.,…, Guisan, A. (2012) High host-plant nitrogen content: a prerequisite for the evolution of ant-caterpillar mutualism? Journal of Evolutionary Biology, 25,1658-1666. doi: 10.1111/j.1420-9101.2012.02555.x

Pennington, T.D. (1997) The Genus *Inga*. The Royal Botanic Gardens, Kew, London, UK.

Pierce, N. (1985) Lycaenid butterflies and ants: Selection for nitrogen-fixation and other protein-rich food plants. The American Naturalist, 125, 888-895. doi: 10.1086/284387

Rehr, S.S., Janzen, D.H. & Feeny, P.P. (1973) L-Dopa in legume seeds: a chemical barrier to attack. Science 181: 81–82. doi: 10.1126/science.181.4094.81

Revell, L.J. (2018) Phytools: R package for phylogenetic tools for comparative biology (and other things). Available at: https://cran.r-project.org/web/packages/phytools/phytools.pdf. Accessed December 22, 2018.

Rhoades, D.F. (1979) Evolution of plant chemical defense against herbivores. In G.A. Rosenthal G.A. & D.H. Janzen (Eds.), *Herbivores: their interaction with secondary plant metabolites* (pp. 3-54). New York, NY: Academic Press.

Richardson, J.E., Pennington, R.T., Pennington, T.D., & Hollingsworth, P.M. (2001) Rapid diversification of a species-rich genus of neotropical rain forest trees. Science, 293, 2242–2245. doi:10.1126/science.1061421

Ryan, S., Cane, K., DeBoer, K., Sinclair, S., Brimblecombe, R., & Hamill, J. (2012) Structure and expression of the quinolinate phosphoribosyltransferase (QPT) gene family in *Nicotiana*. Plant Sciences, 188, 102-110. doi: 10.1016/j.plantsci.2012.02.008

Salazar, D., Jaramillo, M.A., & Marquis, R.J. (2016a) Chemical similarity and local community assembly in the species rich tropical genus *Piper*. Ecology, 97, 3176-3183. doi:10.1002/ecy.1536

Salazar, D., Jaramillo, A., & Marquis, R. J. (2016b). The impact of plant chemical diversity on plant-herbivore interactions at the community level. Oecologia, 181, 1199–1208. doi: 10.1007/s00442-016-3629-y

Salazar, D., Lokvam, J. Mesones, I. Vásquez, M., Milagros, K., de Valpine, P., & Fine, V.A. (2018) Origin and maintenance of chemical diversity in a species-rich tropical tree lineage. Nature Ecology & Evolution, 2, 983–990. doi: 10.1038/s41559-018-0552-0

Schenck, C.A., Holland, C.K., Schneider, M.R., Men, Y., Lee, S.G., Jez, J.M., … Maeda, H.A., 2017. Molecular basis of the evolution of alternative tyrosine biosynthetic routes in plants. Nature Chemical Biology, 13, 1029-1035. doi: 10.1038/nchembio.2414

Schenck, C.A., & Maeda, H.A. (2018) Tyrosine biosynthesis, metabolism, and catabolism in plants. Phytochemistry 149, 82-102. doi: 10.1016/j.phytochem.2018.02.003

Sedio, B.E. (2013) Trait evolution and species coexistence in the hyperdiverse tropical forest tree genus *Psychotria*, PhD dissertation, Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan, USA

Sedio, B.E., Wright, S.J., & Dick, C.W. (2012) Trait evolution and the coexistence of a species swarm in the tropical forest understorey. Journal of Ecology 100, 1183-1193. doi:10.1111/j.1365-2745.2012.01993.x

Sterkel, M., Perdomo, H.D., Guizzo, M.G., Barletta, A.B.F., Nunes, R.D., Dias, F.A., Sorgine, M.H.F., Oliveira, P.L. (2016) Tyrosine Detoxification Is an Essential Trait in the Life History of Blood-Feeding Arthropods. Current Biology 26:288-2193. doi: 10.1016/j.cub.2016.06.025

Valencia, R., Condit, R., Foster, R.B., Romoleroux, K., Munoz, G.V., Svenning, J-C., … Balsev, H. (2004) Yasuní forest dynamics plot, Ecuador. In E.C Losos & E.G. Leigh, E.G. (Eds.) *Tropical forest diversity and dynamism: Findings from a large-scale plot network* (pp 609-620). Chicago, IL: University of Chicago Press.

Vavricka, CJ, Han, Q, Mehere, P, Ding H. Christensen, BM & Li, J. 2014. Tyrosine metabolic enzymes from insects and mammals: A comparative perspective. Insect Science 21, 13-19. doi:10.1111/1744-7917.12038

Vitousek, P.M. (1984). Litterfall, nutrient cycling, and nutrient limitation in tropical forests. Ecology, 65, 285–298. doi: 10.2307/1939481

Vleminckx, J., Salazar, D., Fortunel, C., Mesones, I. Dávila, N. Lokvam, J., … Fine, P.VA. (2018) Divergent secondary metabolites and habitat filtering both contribute to tree species coexistence in the Peruvian Amazon. Frontiers in Plant Science, 9. doi: 10.3389/fpls.2018.00836.

Wiggins, N.L., Forrister, D.L., Endara, M.J., Coley, P.D., & Kursar, T.A. (2016) Quantitative and qualitative shifts in defensive metabolites define chemical defense investment during leaf development in *Inga*, a genus of tropical trees. Ecology and Evolution 6, 478-492. doi:10.1002/ece3.1896

Wink, M. 2016. Evolution of secondary plant metabolism. In: eLS.Chichester, John Wiley and Sons. doi: 10.1002/9780470015902.a0001922.pub3
Wright JS (2002) Plant diversity in tropical forests: A review of mechanisms of species coexistence. Oecologia 130, 1-14. doi:10.1007/s004420100809.1.